

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/87711/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Beaman, C. Philip and Jones, Dylan Marc ORCID: <https://orcid.org/0000-0001-8783-5542> 2016. The item versus the object in memory: on the implausibility of overwriting as a mechanism for forgetting in short-term memory. *Frontiers in Psychology* 7 , 341. 10.3389/fpsyg.2016.00341 file

Publishers page: <http://dx.doi.org/10.3389/fpsyg.2016.00341>  
<<http://dx.doi.org/10.3389/fpsyg.2016.00341>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.





# The Item versus the Object in Memory: On the Implausibility of Overwriting As a Mechanism for Forgetting in Short-Term Memory

C. Philip Beaman<sup>1\*</sup> and Dylan M. Jones<sup>2</sup>

<sup>1</sup> Centre for Cognition Research, School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK,

<sup>2</sup> School of Psychology, Cardiff University, Cardiff, UK

## OPEN ACCESS

### Edited by:

Jerker Rönnerberg,  
Linköping University, Sweden

### Reviewed by:

Emily M. Elliott,  
Louisiana State University, USA  
Lars Nyberg,  
Umeå University, Sweden

### \*Correspondence:

C. Philip Beaman  
c.p.beaman@reading.ac.uk

### Specialty section:

This article was submitted to  
Auditory Cognitive Neuroscience,  
a section of the journal  
Frontiers in Psychology

**Received:** 11 December 2015

**Accepted:** 23 February 2016

**Published:** 10 March 2016

### Citation:

Beaman CP and Jones DM (2016)  
The Item versus the Object  
in Memory: On the Implausibility  
of Overwriting As a Mechanism  
for Forgetting in Short-Term Memory.  
Front. Psychol. 7:341.  
doi: 10.3389/fpsyg.2016.00341

The nature of forgetting in short-term memory remains a disputed topic, with much debate focussed upon whether decay plays a fundamental role (Berman et al., 2009; Altmann and Schunn, 2012; Barrouillet et al., 2012; Neath and Brown, 2012; Oberauer and Lewandowsky, 2013; Ricker et al., 2014) but much less focus on other plausible mechanisms. One such mechanism of long-standing in auditory memory is overwriting (e.g., Crowder and Morton, 1969) in which some aspects of a representation are “overwritten” and rendered inaccessible by the subsequent presentation of a further item. Here, we review the evidence for different forms of overwriting (at the feature and item levels) and examine the plausibility of this mechanism both as a form of auditory memory and when viewed in the context of a larger hearing, speech and language understanding system.

**Keywords:** auditory cognition, short-term memory, memory, forgetting, auditory scene analysis

Like many cognitive capabilities, language is grounded in memory. A failure to appreciate what has just gone drastically limits the ability to comprehend the present and any capacity to anticipate the future. Both long-term memory (for semantics and other lexical and world knowledge) and short-term memory (a record of the immediate past) are implicated in this process. In the current paper, a particular focus is placed upon the relationship between memory and language reception (e.g., hearing) rather than production (e.g., speaking). Although the latter is clearly of importance – both as an aspect of language in which memory must play its part and as a means by which (via overt or sub-vocal rehearsal) information is maintained in short- or longer-term memory ( Craik and Watkins, 1973; Ward et al., 2003; Ward and Tan, 2004; Taylor et al., 2015) – a focus on the nature of the auditory-perceptual input suggests constraints on how any system accepting such input must be configured.

A key feature of the classical short-term memory (STM) research program is the importance of serial order (Lashley, 1951; Conrad, 1960; Murdock, 1968, 1983; Lewandowsky and Murdock, 1989; Henson, 1998; Brown et al., 2000; Botvinick and Plaut, 2006; Burgess and Hitch, 2006). Words, letters, or digits are presented sequentially and participants required to recall the items in the order in which they were presented. Short-term memory tasks are usually deliberately designed so that the associations to be held across multiple items (words, digits, letters) are arbitrary. Performance in such tasks is framed in terms of its proximity to verbatim recall of all the items, namely the correct item in its position at presentation. Implicitly, short-term memory theorists make the assumption that the “item” (the word, letter, or digit of interest)—rather than the relationship

between items—is the most meaningful unit of analysis. Moreover, identification of the item at recall is the basis for correct scoring in the memory test. Despite known problems with identifying the “item” in anything other than a logically circular way (Miller, 1956) such an approach is defensible in cases where the items are well-known and taken from a small, circumscribed set (e.g., digits) and where recollection of an item at a time collapses into a requirement to select the most likely candidate given the degraded or incomplete information available (Nairne, 1990). This contrasts starkly with the situation in most everyday language, in which structured, non-arbitrary relationships are available between individual elements represented at multiple levels (phonotactic, syntactic, semantic, pragmatic) and the identification of a single “item” is neither necessary nor sufficient to comprehend the meaning of the sequence.

By considering veridical recall of arbitrary items rather than the relationships between them, much of interest is lost with regards to later analyses. A key component of perception is in organizing as well as registering information and of interest is whether, in registering and organizing the stimuli prior to retrieval, the perceptual system represents them in a way that harmonizes with the retrieval requirements in standard short-term memory tasks. Given the emphasis on the iterative retrieval of items across the to-be-remembered sequence, does the perceptual system, for instance, cluster items at a supra-item level in such a way as to aid or to hamper efficient retrieval? In other words, does perception result in ‘items’ corresponding exactly to items specified in terms of the linguistic taxonomy (such as single syllables, words, or digits) on which the sequence is nominally based? In the event of supra-item organization, how are items grouped or transformed? Is there grouping of adjacent elements (as with chunking, classically) or are non-adjacent items organized into a greater whole? Within item-focussed approaches to short-term memory—ones that assume recall is a product of an aggregation of elemental actions—forgetting may be explained by the ‘overwriting’ of items by subsequent events. If supra-segmental organization occurs, is overwriting still a plausible mechanism?

Here, we explore how the registration of events in memory reflects auditory input and, in particular, the organizational processes that are at play. On the basis of key phenomena in auditory perception we consider potential implications for the structure of short-term memory and, in particular, the nature of forgetting.

## THE “STANDARD” MODEL OF MEMORY

The modal model of memory, informed by neuropsychological case data, has always assumed a functional and structural distinction between short-term and long-term memory, with the former fed by largely unspecified perceptual input processes, frequently depicted as a buffer storage system (Shallice and Cooper, 2010). In long-term memory, where the notion of memory as a reliable, veridical system has long since been dismissed and a reconstructive account of recall is generally accepted (Bartlett, 1932), suppression, inhibition and blocking of

the memory trace have all been discussed as possible explanations of forgetting (for example in the context of the misinformation effect in eyewitness memory). In contrast, discussions of short-term and sensory memory have been less open to the idea of memory distortion as normal and recall as a reconstructive activity. In consequence, processes that highlight deterioration of the representation such as decay and overwriting (respectively) have predominated as mechanisms for forgetting and active supra-segmental organizational processes (such as grouping into objects), which may equally hamper recall when they are inconsistent with retrieval requirements, have been largely ignored.

Much has already been written both critiquing the evidence for decay (e.g., Neath and Nairne, 1995; Nairne, 2002; Lewandowsky and Oberauer, 2008, 2009, 2015; Lewandowsky et al., 2008; Oberauer and Lewandowsky, 2008, 2013; Brown and Lewandowsky, 2010; Neath and Brown, 2012) and defending the concept (Altmann and Gray, 2002; Portrat et al., 2008; Altmann, 2009; Barrouillet et al., 2011; Altmann and Schunn, 2012) so, rather than repeating now-familiar arguments about decay versus some other (often unspecified)<sup>1</sup> form of interference as the source of forgetting (see Ricker et al., 2014, for a review), here we will specifically consider interference by overwriting as it appears from the perspective of auditory perception and the organization of the auditory environment.

The introduction of overwriting or displacement as a key determinant of forgetting over the short-term can be traced back to early studies of auditory sensory memory. Classically, a restricted-capacity acoustic sensory memory trace, overwritten by subsequent auditory events (Crowder and Morton, 1969), is available to supplement end-of-sequence recall otherwise only supported by “post-categorical” short-term memory systems dedicated to verbal memory but otherwise blind to sensory modality or the perceptual origins of the memoranda. This venerable account is nonetheless still extant and has been incorporated into more recent formulations of the contribution of sensory memory to immediate recall of auditory-verbal material (e.g., Page and Norris, 1998, and Burgess and Hitch, 1999, both make reference to an auditory input buffer overwritten by subsequent data). Latterly, other formulations of short-term memory have also utilized overwriting as a means of implementing interference and hence forgetting. For example, in providing a framework for short-term memory that eschews decay as a concept, Nairne (1990, 2002), Neath and Nairne (1995), Neath (2000), Oberauer and Kliegl (2006), Oberauer and Lange (2008), Lewandowsky et al. (2009), and Oberauer (2009) explicitly replace decay with overwriting as an explanatory concept. For memory of specifically auditory origin, therefore, three claims have been made:

- (1) An auditory sensory store is overwritten, an item at a time, during encoding (e.g., Crowder and Morton, 1969; Page

<sup>1</sup>In fact, Lewandowsky et al. (2009, box 4) postulate at least four possible alternatives to trace decay and Ricker et al. (2014, Table 1) suggest five possibilities, all of which – as with decay itself – may be implemented in different ways (e.g., Lewandowsky and Farrell, 2011).



and Norris, 1998; Burgess and Hitch, 1999; Mercer and McKeown, 2010a).

- (2) Both modality specific (sensory) and modality-independent (post-categorical, phonological) features are overwritten during encoding (e.g., Neath and Nairne, 1995; Neath, 2000).
- (3) Features are overwritten via neural competition removing the availability of feature-units at a post-encoding phase (Oberauer and Kliegl, 2006; Oberauer and Lange, 2008; Oberauer, 2009).

It is interesting that the “precategorical” acoustic nature of the auditory sensory store (Crowder and Morton, 1969) arose because of prior theoretical commitments to a model of word recognition—the logogen model—which assumed a single system for recognizing both written and spoken words (Morton, 1964, 1969). Subsequent to this, changes to the logogen model (Morton, 1979) removed this theoretical constraint and introduced separate auditory and visual input logogens so that the idea that overwriting occurred at an early processing stage was retained even though the original *a priori* reasons for assuming that overwriting occurred prior to word-identification had vanished. The Crowder and Morton (1969) view is, despite its commitment to a pre-categorical (presumably continuous) representational format a classically item-based data buffer system of the first-in, first-out variety. Their approach can be contrasted with the forms of overwriting implemented in models by Nairne (1990) and Lange and Oberauer (2005).

In Nairne’s (1990) feature model of immediate memory, individual items are represented as vectors of features, which may represent modality-specific or modality-independent information. The eponymous features were speculatively identified with patterns of neural firing by Beaman et al. (2008) and, although their exact nature and status has never been formally defined, it is at this level that overwriting operates within the model. Feature overwriting works by an incoming item deleting identical features already held as part of the representation of immediately preceding item. For example, if the third feature of item  $n+1$  of a sequence takes the same value as item  $n$  of the same sequence then the item-level representation of  $n$  is denuded of this feature, the representation becomes degraded as a consequence and  $n$  is henceforth less likely to be correctly recalled when cued to do so at some point in the future.

In contrast, the version of overwriting put forward by Lange and Oberauer (2005) and Oberauer and Kliegl (2006) interference is not limited to the preceding item. Like the approach of Nairne (1990), the model is once again feature-based; in this instance, however, different items are represented as patterns of activation across a subset of the features (“feature units”) available system-wide and representations compete for access to their constituent feature units. Where a given representation loses this competition, the feature unit is captured by that competitor and is not available as part of the item representation of the “losing” representation. In this way a particular representation is degraded, thus impeding recall. The neural competition for features is framed in terms of synchronized firing of neurons as a mechanism of binding together the features that

belong to the representation of an item (Raffone and Wolters, 2001). Feature units possessing features belonging to the same representation fire synchronously, whereas units belonging to different representations fire out of synchrony.

The principal difficulty with overwriting as the sole, or key, determinant of failure to recall in these or any other accounts is that while many studies have reported greater interference when irrelevant information (e.g., from a secondary task; Lange and Oberauer, 2005) is related to the memoranda, or when the list items are themselves similar along a specific dimension (e.g., the phonological similarity effect; Conrad, 1964; Conrad and Hull, 1964; Baddeley, 1966) other studies have shown the opposite. Overwriting in the three accounts given above assumes that interference occurs between similar items or items with similar features – acoustic items displace earlier acoustic items in a precategorical store (Crowder and Morton, 1969) or features are overwritten if they are shared between successive items (Nairne, 1990) or if they are supported by common feature units (Lange and Oberauer, 2005). These assumptions readily account for data in which interference is observed at recall between items that are similar along one or more crucial dimensions. However, Mercer and McKeown (2010a,b) found that complex tones were more accurately identified in a same-different task when followed by distractors containing *novel* frequencies – those frequencies not present in the target – when compared to a condition in which the distractors shared frequencies with the target. This pattern of results is directly contrary to that which would naturally occur if similarity-based overwriting was in operation.

Interestingly, Mercer and McKeown (2010b) also concluded in favor of an overwriting account – but in their model, directly contrary to assumptions made by other theorists about overwriting, “interference is principally caused by tones that include *novel* features since these will be most potent in “overwriting” the contents of the auditory spectral short-term memory buffer” (Mercer and McKeown, 2010b, p. 1258, emphasis added). In other words, this model assumes overwriting by items which are representationally distinct from the preceding input, rather than by items which share features with earlier items. Whether overwriting is assumed to occur amongst similar or dissimilar items/features is, of course, an *a priori* decision for any theorist attempting to construct a model (Lewandowsky and Farrell, 2011) but it is unlikely that similar items would be overwritten in some cognitive systems and dissimilar items overwritten elsewhere. To allow that closely related cognitive and perceptual subsystems work on diametrically opposed principles is, at best, un-parsimonious and contrary to Occam’s Razor. If overwriting is to be accepted then a consistent set of rules should apply (Surprenant and Neath, 2009). Nor is the study by Mercer and McKeown (2010b) (which involved fairly “low-level” and non-verbal acoustic stimuli) unique in its findings. An earlier study by Nairne and Kelley (1999) showed that the phonological similarity effect observed with verbal stimuli is reversed after relatively brief periods of distraction, resulting in better performance in a serial order reconstruction test for phonologically similar lists than for phonologically dissimilar lists. If overwriting is seen as necessary to account for forgetting

caused by interference effects between similar items, then reversing these similarity effects casts doubt upon the need for overwriting.

Finally, task requirements—which are unlikely to directly influence low-level processes such as overwriting/displacement of patterns of neural firing or competition for neural feature units—also play a substantial role in similarity effects for which overwriting is offered as an explanatory mechanism. Despite numerous documented similarities between immediate free and serial recall (Beaman and Jones, 1998; Bhatah et al., 2006, 2008; Ward et al., 2010; Grenfell-Essam and Ward, 2012; Grenfell-Essam et al., 2013; Spurgeon et al., 2014) similarity effects within the to-be-recalled list—supposedly reflecting the impact of item-representations degraded by direct over-writing (Nairne, 1990) or competition for specific feature units (Oberauer and Kliegl, 2006)—depress performance on immediate serial recall tasks but enhance performance in free recall (Fournet et al., 2003). Once again it is difficult to reconcile such findings with a low-level, item-based and automatic overwriting interference process without appealing to a higher-level activity that negates, and more than negates, the negative effect of similarity-based overwriting. To account for the reversal of the phonological similarity effect, Nairne and Kelley (1999) proposed that a period of distraction allows phonological similarity to be used as a cue to select candidate items for serial reconstruction of order (e.g., any correct item must share a rime with all other items) and similar suggestions are equally applicable to free, or item, recall situations (e.g., Watkins et al., 1974; Saint-Aubin and Poirier, 1999). However, such accounts are necessarily *post hoc* and—if overwriting occurs—strategies such as these must be sufficiently ubiquitous and powerful enough to not only negate but *reverse* the similarity effects otherwise observed. Exigencies of space mean that the interesting issue of retrieval mode and streaming cannot here be addressed fully but we note that free recall is in part controlled by strategic retrieval factors so that we may expect effects such as those of similarity to be different dependent upon the mode of recall and, critically, scoring technique employed. A stream of similar-sounding items will necessarily lose order cues relative to a dissimilar stream up until the point that items become so dissimilar that stream coherence is lost (Jones et al., 1999). There are no such necessary consequences for retrieval of individual items so scoring criteria at test are crucial in the appearance and form of similarity effects.

## AUDITORY SCENE ANALYSIS: SOME PRELIMINARIES

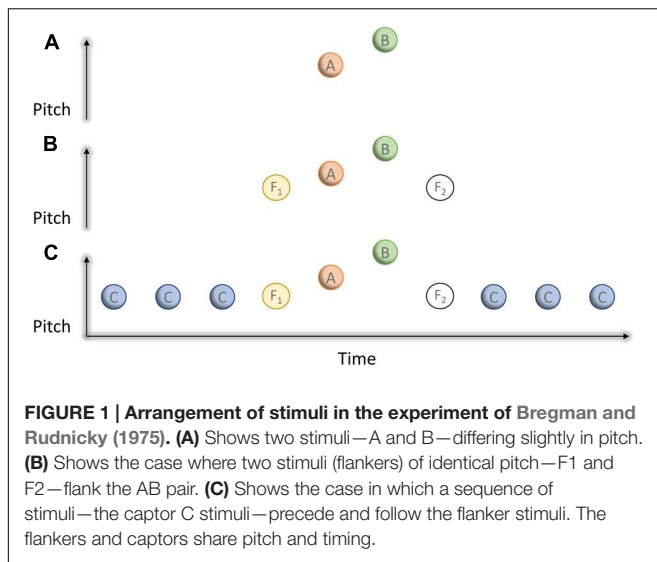
If a structural account of forgetting is set aside, what remains? Perceptual organization has profound consequences not just for the coherence of our experience of the world but also for the accessibility of information contained within it. Perception itself is directly linked to memory, as, for example, the perception of loudness is determined by a temporal integration of acoustic power; the perceived loudness of a burst of white noise depends upon its duration (Scharf, 1978) demonstrating that perception is reliant upon memory in a manner which renders the simple

idea that incoming stimuli “automatically” overwrite pre-existing representations problematic. There is a mass of evidence showing powerful effects of perceptual organization and, as with vision, it is useful to think of auditory perception in terms of objects. So, despite being intrinsically evanescent in a way that the visual world is often not, successive events are assembled into temporally extended objects in a way that allows several “streams” of information to co-exist. Note that this is immediately different from the situation assumed within most models of verbal STM, which concentrate upon memory for a single list, and require further work to allow simultaneous representation of multiple lists or streams within the same representational space. Generally, the rules of organization follow Gestalt principles that are based on the physical attributes of the stimuli: proximity, similarity, closure, symmetry, common fate, continuity, among others. So, auditory perception is an active process that partitions the auditory world into auditory objects or streams, a process known generically as *auditory scene analysis* (Bregman, 1990). It is difficult to overstate the importance that these forces of organization have on what may be retrieved from an auditory scene, even when the scene comprises a few simple stimuli.

Necessarily, stream formation involves memory. A succession of individual stimuli achieves stream quality by a process that depend on not just a single but many preceding stimuli, a process that requires storage. Streams take time to form and less compelling streams can vacillate and break down. In everyday environments, scene analysis typically results in several simultaneous streams, such as the instruments of a rock band or orchestra, or indeed a domestic scene of refrigerator noise, radio and conversation. Also, the principles by which this is achieved are embodied in musical polyphony: the rules of composition—though in a non-acoustic language—allow a composer to generate an intelligible and coherent rendition of harmonic and melodic intent.

So, the logic adopted here is that auditory memory is intimately connected to auditory perception and that in turn the study of auditory perception suggests ways in which auditory-verbal memory is organized. Furthermore, we know that this organization is not veridical, in as much as it does not faithfully represent an item-by-item sequence, free of item clusters. As we will see later, the item-clusters produced by auditory perception are very much richer and more diverse than those considered by current models of verbal STM.

It is useful to consider specific instances, using some very simple non-verbal stimuli, of how perceptual organization of sound brings about changes to perception before returning to the case of verbal memory. The first example shows how the context in which stimuli appear works to shape what we may know of them. Take the very simple case of two short tones, A and B, the same in every respect except that they are a semi-tone apart, presented in quick succession (see **Figure 1**). When faced with the task of reporting the order as being high-low or low-high, most listeners find they can make the discrimination quite easily. However, if flankers ( $F_1$  and  $F_2$ )—sharing almost the same pitch and tempo as A and B (see again **Figure 1**)—are inserted either side of them then we observe a dramatic reduction in the capacity to report the order of A and B. How might this come about?



One way to think in terms of overwriting and to suppose that the second flanker (or indeed both flankers) somehow interfere with the representation of A and B, making their comparison less easy. Another way is frame the change in context in terms of object formation. Whilst presented as a pair, A and B formed a single object and at the same time constituted its boundaries. Adding the flankers created a new object and new boundaries, with A and B now constituting its innards, so that now the order information contained in A and B becomes more difficult to address. This is a familiar situation in STM where current recall of the items and order of the first and last few items gives rise to primacy and recency effects, with recall of items in the correct order very much worse toward the center of the list.

A simple further addition to this auditory scene shows how implausible the overwriting explanation turns out to be in the case of simple tones. If we add a further two stimuli ( $C_1$  and  $C_2$  in **Figure 1**) either side and sharing both pitch and tempo with the flankers then we witness a remarkable transformation: if we now ask a listener to judge the pitch order of A and B, close to full efficiency (that is, the level of performance when A and B are presented in isolation) is restored. Clearly, according to the overwriting view (and indeed, most interference theories of forgetting) adding more stimuli should – if anything – produce more overwriting, not less. However, the outcome of adding the C stimuli is readily understood in terms of auditory scene analysis. The C stimuli act as ‘captors,’ that is, by virtue of their greater similarity to the F stimuli than to the AB pair, two objects are formed; the one: CCCF<sub>1</sub>F<sub>2</sub>CCC, the other: AB. The flankers are captured, releasing AB to become a separate, and therefore an independently addressable, entity thereby restoring memory for the order of A and B.

This setting shows several remarkable qualities of auditory scene analysis with a number of important implications for the way we understand memory. The first and most profound is that the future shapes the past: perception is retroactive. What follows from this has great relevance to our current discussion about the plausibility of overwriting as an explanatory construct.

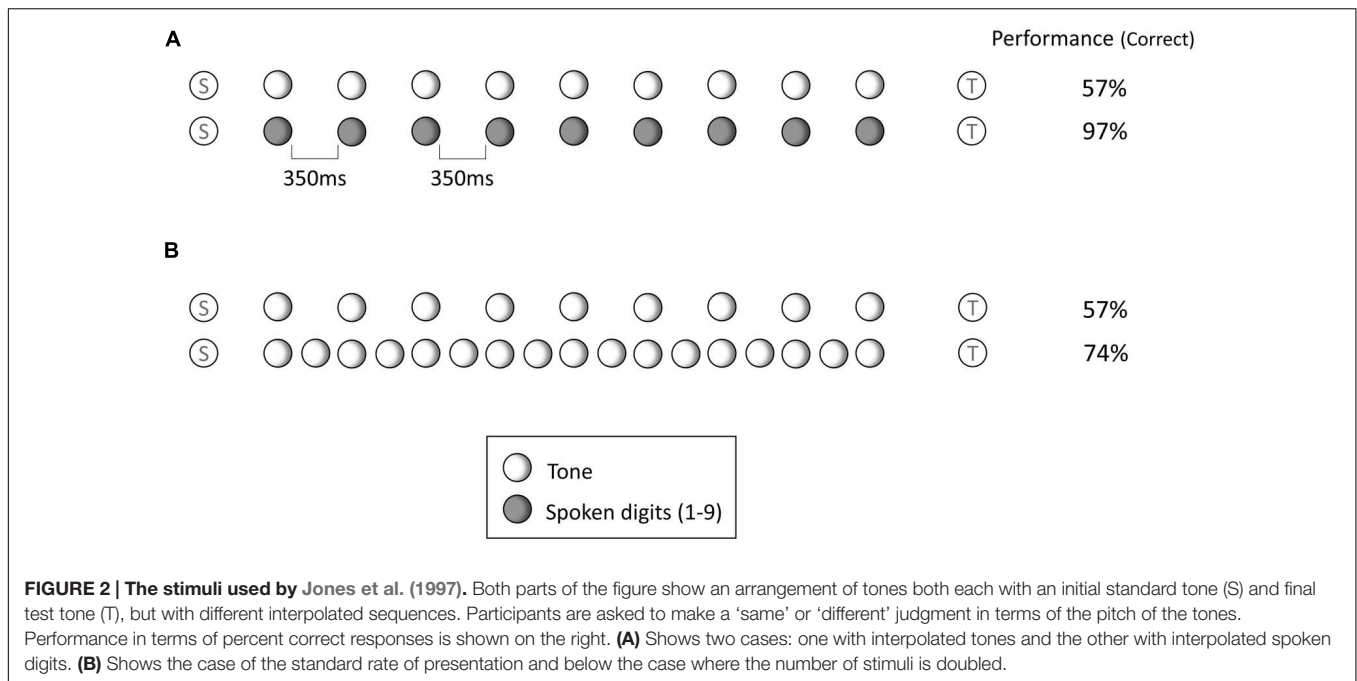
Critically, the perceptibility of AB is only decided when both F<sub>2</sub> and CCC are presented, but even then both F<sub>2</sub>CCC and AB are distinguishable only with reference to F<sub>1</sub> and CCC. The first point that follows from this is that it is important therefore to think in terms of the emergent properties of the stimulus ensemble (the object), not merely as an aggregation of the properties of individual stimuli. The second point is that items need not be temporally adjacent in order to form into objects.

A second illustration lends weight to the first while at the same time addressing the natural skepticism that such a simple setting involving the mere ‘perception’ of tones A and B could have more general repercussions for more complex settings that we think as being characteristic of the study of ‘memory.’ Here again, the listener is asked to compare two tones but this time asked to make the judgment about whether they are the same or different in pitch (Jones et al., 1997).

**Figure 2** shows the arrangement of stimuli used by Jones, Macken, and Harries (following, for example, Deutsch, 1972, 1978a,b; Semal and Demany, 1991, 1993; Starr and Pitt, 1997; Mathias and von Kriegstein, 2014). First, a standard stimulus—a tone—is followed either by a blank interval or a filled interval and then, some seconds later, by a comparison stimulus: another tone. The listener is asked to ignore stimuli that come between the standard and comparison tones in making their judgment.

The key variable of interest is the content of the interval and its effects on the accuracy of the comparison judgment. Having a sequence of tones in the interval similar in pitch and timbre to the standard and comparison (see **Figure 2**) has a dramatic effect of reducing the accuracy of the same-different judgment. If, instead of having tones, we have speech stimuli (say a sequence of words), comparison judgment improves considerably, to a level that is close to when there are no interpolated stimuli. This result is conventionally interpreted in an overwriting framework: memory for the standard is compromised by similar stimuli interpolated between it and the comparison (e.g., Semal and Demany, 1991, 1993; Mathias and von Kriegstein, 2014). However, another of the conditions in the study of Jones et al. (1997) makes this interpretation implausible. If the number of interpolated tones is doubled then any reasonable interpretation the overwriting account suggests that performance cannot improve and should, in fact, deteriorate. If overwriting interferes only with the immediately preceding item (as with Nairne, 1990) then the level of interference remains the same, although the increase in the number of sources competing for consideration at recall could still negatively affect overall performance. If overwriting is not restricted to immediately preceding items (as with Lange and Oberauer, 2005) then performance should deteriorate, and appreciably so given the rise in number of interfering sources. In the event, the opposite turns out to be true; performance improves significantly.

If we construe the setting in terms of auditory scene analysis, this last result is entirely intelligible. In object terms, the proximity of the standard to the interpolated tones and the similarity of their physical character (sharing tone-like qualities), along with its shared timing, increases the likelihood that it will be incorporated with them into an object, thereby reducing its identity as a separate entity. When the interpolated material is



speech, of course this tendency will be much less likely. Doubling the number of interpolated tone stimuli is likely to produce an outcome similar to that seen with interpolated speech stimuli: by virtue of shared timing (in addition to shared pitch and timbre) the interpolated stimuli will in this case form an object separate from the standard. The judgment of similarity is once again based on two stimuli distinct from the interpolated stimuli: the scene comprises three objects, a standard, a distinct interpolated stream, and the comparison.

Streaming thus produces important consequences for our judgment of the plausibility of overwriting as an explanatory mechanism and for hearing and memory. The context in which stimuli appear has powerful repercussions for what we can retrieve of stimuli. As we shall go on to consider, the fact that auditory stimuli appear in chronological order does not mean that that access to temporally adjacent stimuli is guaranteed. So, for instance, if we present a sequence in alternating male-female voices ( $M_1F_2M_3F_4M_5F_6$ ), two streams are formed ( $M_1M_3M_5$  and  $F_2F_4F_6$ ) a situation that contrasts with a single (e.g., male) stream:  $M_1M_2M_3M_4M_5M_6$ . By forming two distinct streams it will become more difficult to retrieve chronologically adjacent stimuli (e.g.,  $M_1F_2M_3F_4$  will be harder to retrieve than  $M_1M_2M_3M_4$ ), but easier to retrieve stream-adjacent (and chronologically non-adjacent) stimuli (e.g.,  $M_1M_3M_5$  will be easier to retrieve in the alternating voices case) if cued to retrieve the utterances in strict temporal order of their occurrence. Notice that—as suggested earlier—the stream can contain non-adjacent elements. This contrasts with the typical interpretations of ‘chunking’ (and also “grouping”) that invariably refer to an aggregation of temporally adjacent elements. Auditory scene analysis shows that even quite remote elements may be assembled into an organized whole. This is why scene analysis and chunking are slightly different mechanisms and why it is important to consider remote elements

in any scene analysis (see Jones, 1993, 1999; Jones et al., 1996; for extended discussions). This relates to the question of overwriting because temporally remote and non-adjacent items can have a greater effect upon memory for any given target item than does the immediately subsequent item, a result which is inconsistent with at least two forms of overwriting (Crowder and Morton, 1969; Nairne, 1990)

Perhaps the simplest and most telling prediction from the overwriting hypothesis is that sequences with fewer shared features should be easier to retrieve than those with many shared features. This follows from such ideas as the relative distinctiveness principle, the suggestion that an item (or series of items) perceived to be discriminable on some dimension(s) from its fellows is easier to recall by virtue of psychological distinctiveness (a principle which is consistent with overwriting as an underlying mechanism, although other mechanisms may produce such an outcome; Surprenant and Neath, 2009; Neath, 2010). Evidence already reviewed indicates that this is not always the case, and further data indicate that streaming may be a useful concept in explaining outcomes that run contrary to this principle.

Very distinct non-speech sounds, when presented quickly in a sequence are easy to recognize, so that listeners can judge they are present but are less able to indicate the order in which they appeared. So, if a sequence of very different sounds—a high-pitched tone, a hiss, a low-pitched tone and a buzz—are heard in a repeating cycle, listeners are able to name each of the sounds. However, they cannot report their order correctly, even if the period of listening is extended indefinitely (Warren et al., 1969; see also Jones et al., 1999). However, if a sequence of four spoken digits—spoken in the same voice—is presented under the same conditions, the order can be readily reported. The key difference between these two settings is in the level of commonality in



acoustic content: acoustically the digits form a variation on a common ground and so quickly form a stream, but for the non-voice sounds each element constitutes a separate entity and streaming is less easy to achieve.<sup>2</sup> Consider how such a situation would be addressed by Nairne's (1990) feature model, in which automatic overwriting forms a large part. The identity of the stimuli themselves would be represented in secondary memory, so the task would simply be to match the correct primary memory representation to the correct secondary memory identity in the correct order. The task would be made difficult by the fact that overwriting would degrade the primary memory representations such that the primary-secondary memory match would become more problematic and, potentially, confused. This confusion would clearly be more prevalent in the situation under which the most overwriting occurred – when the stimuli come from a common source (spoken digits) and share common acoustic and lexical features. These results pose grave difficulties for an overwriting account; distinct sequences should be subject to less overwriting, but the results are diametrically opposite. The explanation comes from stream formation: when the stimuli are perceived as originating from a common source they form a single stream within which order is preserved.

## AUDITORY SCENE ANALYSIS: THE 'SUPRASEGMENTAL' APPROACH APPLIED TO VERBAL MEMORY

In view of the problems outlined earlier, we wish to outline an alternative framework in which retrieval (in both sensory and short-term memory) is constrained by perceptual principles. The primary line of argument we wish to pursue is that the need to maintain a coherent stream of information over time constrains the processes operating within memory and hence automatic and immediate overwriting of an item representation or the features representing an item is not a tenable explanation for forgetting. The auditory scene analysis principles outlined above, however, were introduced with reference to simple auditory stimuli (e.g., tones) and in what follows these are expanded to encompass more traditional verbal memory phenomena.

Within auditory memory, overwriting was originally proposed as an explanation for the interference associated with a post-stimulus suffix (Crowder and Morton, 1969) so we turn first to this phenomenon and possible alternative accounts.

### The Failure of Overwriting: Capturing the Suffix

Classically, the existence of acoustic storage termed the "precategory acoustic store" (PAS; Crowder and Morton, 1969) was assumed to precede "post-categorical" verbal storage (where modality of origin – spoken or written – is irrelevant, a common abstract representation is shared by all stimuli, regardless of input modality). Its existence was inferred from the auditory recency

effect, in which the final item of an auditorily presented list for serial recall is recalled at near-ceiling levels compared to the much smaller recency effect obtained with visually presented lists of the same verbal items. The reason that this has been attributed to a restricted capacity acoustic store is that elsewhere along the list performance on visually and auditorily presented lists is broadly equivalent (but see Beaman, 2002; Macken et al., 2015) and is affected in a similar manner by standard verbal manipulations such as phonological similarity, word-length, and concurrent articulation. The final piece of evidence provided in support of PAS was that the presence of a post-stimulus suffix effectively eliminates this final-item advantage, leading Crowder and Morton (1969) to conclude that the stimulus suffix effect "depends upon selective displacement of information from PAS" (Crowder and Morton, 1969, p. 369).

Crowder and Morton (1969) assumed an item-by-item displacement system rather than feature-based overwriting and one reason for disputing the feature-based interference account of the stimulus suffix effect comes from data showing that stimulus suffixes which are phonemically similar to the memoranda do not necessarily show larger suffix effects (Crowder and Cheng, 1973) and—like the other similarity-based interference effects already reviewed—may also show smaller effects (Carr and Miles, 1997). Another reason for questioning feature-based overwriting comes from studies of *streaming* the suffix. It is well established that the stimulus suffix effect depends at least in part upon the suffix being perceived as originating from the same sources, or stream, as the to-be-recalled list. So, for example, variations in the spatial location, timbre and pitch of the suffix relative to the list reduces the size of the suffix effect whereas similar manipulations varying suffix frequency, emotionality and meaning have no such effect (Morton et al., 1971). Other manipulations varying the "speech-like" qualities of the suffix similarly moderate the size of the suffix effect (Morton et al., 1981). Manipulations of the top-down interpretation of the suffix likewise show that forcing the suffix to be grouped with, or apart from, the list affects the auditory memory interference effect (Crowder, 1971; Frankish and Turner, 1984; Neath et al., 1993). So, for example, ambiguous stimuli, which can be perceived as either speech or non-speech, can be treated as a speech suffix on the basis of labeling them as such (Ayres et al., 1979; Neath et al., 1993). However, other non-speech stimuli do not show a suffix effect unless contextual effects also force them to be perceived as speech (Morton and Chambers, 1976; Ottley et al., 1982). These results show that physically identical stimuli, which bear physically identical relationships to the memoranda, can produce different memory effects depending upon context and expectation. At best, therefore, any interference effect obtained under such circumstances can only be ascribed only in part to the physical overwriting of the memory trace.

Perhaps most intriguingly, the effects of a repeated suffix have also been shown to reduce the disruption observed (Crowder, 1971, 1978; Morton, 1976). With a repeated suffix, the same suffix item is presented multiple times in quick succession and in tempo with the list sequence (as usually also happens with a single suffix). The reduced effect of the suffix when repeated in this way, even though the first presentation of the repeated

<sup>2</sup>The rate of presentation in these studies is fast and this prevents verbal labeling; when the speed of presentation is slowed performance improves but only to a relatively small degree.



suffix is physically identical to the presentation of a single suffix, is difficult to reconcile with an overwriting account based upon physical or feature similarity between successive items since the relationship between the suffix and the list items is equivalent in the two conditions. Critically, a repeated suffix only becomes a repeated suffix at the point of its re-presentation; logically, therefore, automatic overwriting occasioned by the first presentation of the suffix must already have occurred at this point. Data such as these have led to suggestions that the suffix effect might reflect the simultaneous action of overwriting, accounting for the reduced suffix effect still observed, and perceptual grouping, accounting for the difference between single and repeated suffix effects (e.g., Morton, 1976). According to these accounts, the repeated suffix forms a perceptual group apart from the to-be-remembered list whereas the single suffix is perceived as part of this list. It follows from this that the sole cause of the disruption observed in the repeated suffix condition is from overwriting. A single item suffix likewise overwrites the final item but also further depresses memory performance by increasing the functional size of the memory set (the list length) by an extra item (e.g., Nairne, 1990).

These data undermine the importance of overwriting as the source of the suffix memory disruption effect but do not rule out the possibility that overwriting occurs; perhaps it is merely contributing only part of the observed disruption. Later data reported by Nicholls and Jones (2002) are, however, less equivocal. In their experiments, Nicholls and Jones (2002) interleaved a sequence of irrelevant items between the to-be-remembered list items such that the suffix, when presented, was perceptually grouped with, or “captured” by, these irrelevant items. The sequence comprised the item ‘ah,’ which was also used as the suffix in a traditional suffix effect condition (see **Figure 3**). When no-suffix, suffix and captured suffix conditions are compared, it is clear that in the captured suffix condition performance approximates that to the no-suffix condition<sup>3</sup>. In the captured suffix condition the recency effect was fully restored and there was no suffix effect on the final list-item when the suffix was grouped, or streamed, with the sequence of irrelevant items. In contrast, the suffix presented alone continued to produce a suffix effect. Unlike the repeated suffix manipulation which reduced but did not eliminate the suffix effect, these data cannot easily be explained by the joint operation of overwriting and grouping since—in this case—the grouping (or streaming) manipulation removed the suffix effect entirely and hence the need to assume overwriting as the basis of the suffix effect.

Thus, the proposition that auditory-sensory memory is necessarily automatically overwritten is untenable. However, the suffix effect is only a single line of evidence. Recently, doubts about overwriting have been reinforced by findings from a paradigm using alternating voices for each list item and observing the consequences for memory of streams created in this way (Hughes et al., 2009). A suffix presented in a different voice reduces the suffix effect (Morton et al., 1971), consistent with the idea that overwriting depends on similarity between

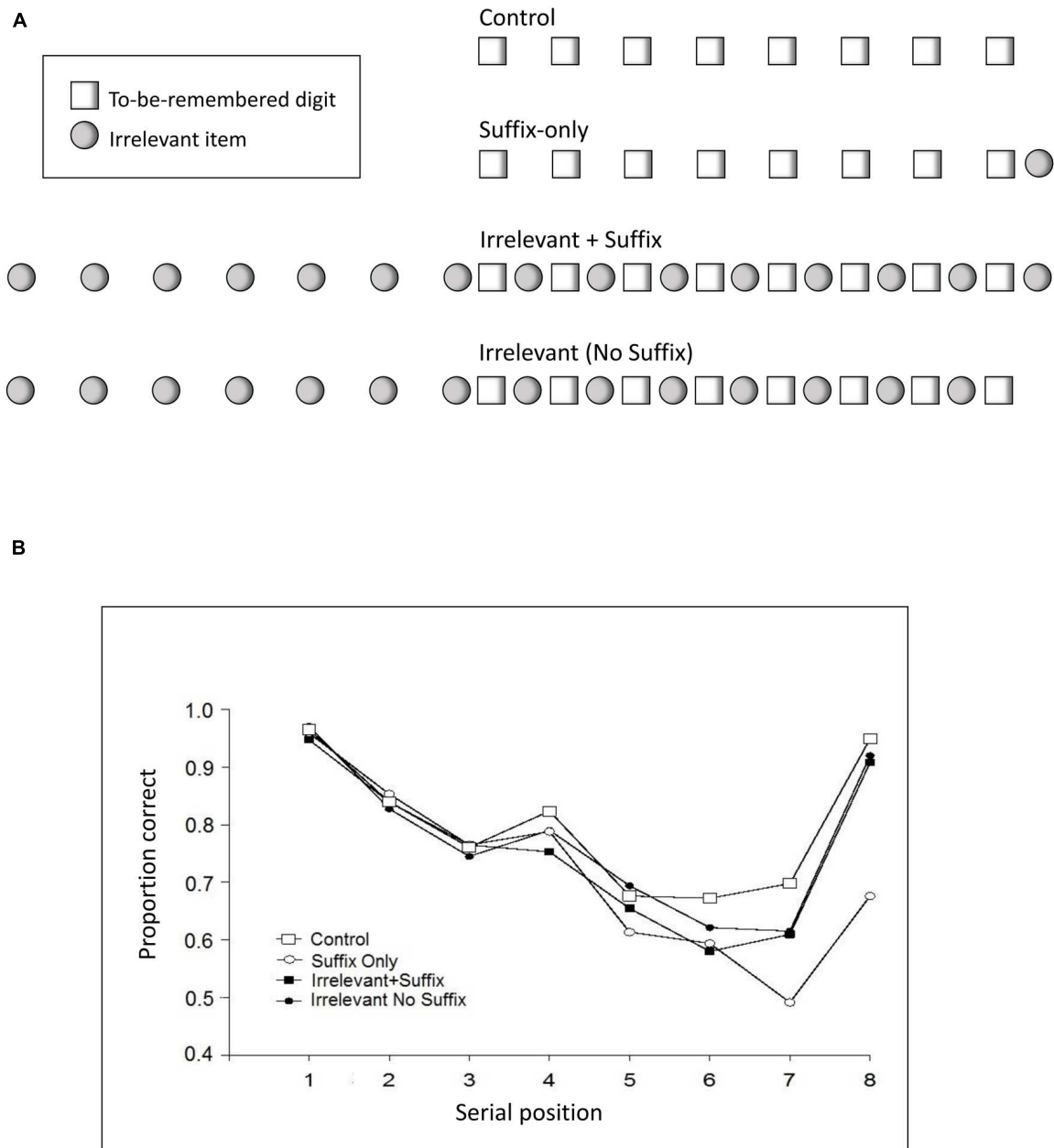
the suffix and the final list item but also consistent with the idea that a different voice suffix is grouped apart from the list items. If overwriting is automatic and based solely upon such physical properties and relationships between successive items, then presenting the to-be-recalled list in alternating voices (e.g., male-female-male and so on), should limit the overwriting observed between successive items compared to the same items presented in a single voice because the feature similarity is reduced by the voice change. Hence, overall recall should be enhanced relative to single-voice presentation. Alternatively, if perceptual organization is important so that items presented in different voices are streamed as coming from distinct sources, then recalling the items in the correct serial order should be harder. As noted in early research on auditory attention, items are preferentially recalled according to the stream or channel from which they are perceived to originate (e.g., Broadbent, 1958; for an extensive discussion see Hughes et al., *in press*) such that if the two voices are perceived as two separate streams then to recall the items in correct serial order requires participants to shift alternately between streams in order to reconstruct the serial order of the list. This extra cognitive requirement imposes a behavioral cost such that a list of alternating voices is not recalled as well as the same items presented in a single voice (Hughes et al., 2009; see **Figure 4**). Again this talker-variability effect calls into question the predominance of overwriting, which would predict the opposite pattern of results.

## Time, Space and Voice-Based Grouping Effects

The talker-variability effect, together with the different-voice suffix effect, supports the assumption that lists presented in different voices are perceptually grouped apart and that this influences the appearance of memory phenomena. Such assumptions find further support from early work on auditory attention (Broadbent, 1958) together with current theories of low-level auditory perception, within which auditory stream segregation (Bregman, 1990) plays a central role. One further line of evidence, however, serves to emphasize the relationship between perceptual organization and what seem superficially to be wholly mnemonic processes (suffix and talker-variability effects).

Work on grouping within auditory memory by Frankish and Turner (1984), Frankish (1985, 1989, 1995) directly examines the effect of perceptual grouping on subsequent recall. In a series of experiments, Frankish (1985, 1989, 1995) demonstrated that coherent groups can be formed within lists presented for immediate serial recall. These groups are defined by boundaries that exhibit the same, or similar, primacy and recency effects at recall as the longer lists of which they form a part. For example in a control (ungrouped) list, recency occurs only at the end of the list. However, in a 9-item list which is organized into three groups of three items each—for example by a delay in presentation between items 3 and 4 and between items 6 and 7—recency is seen for the final item of group 1 (at serial position 3, which must therefore be relatively immune to the suffix effects of

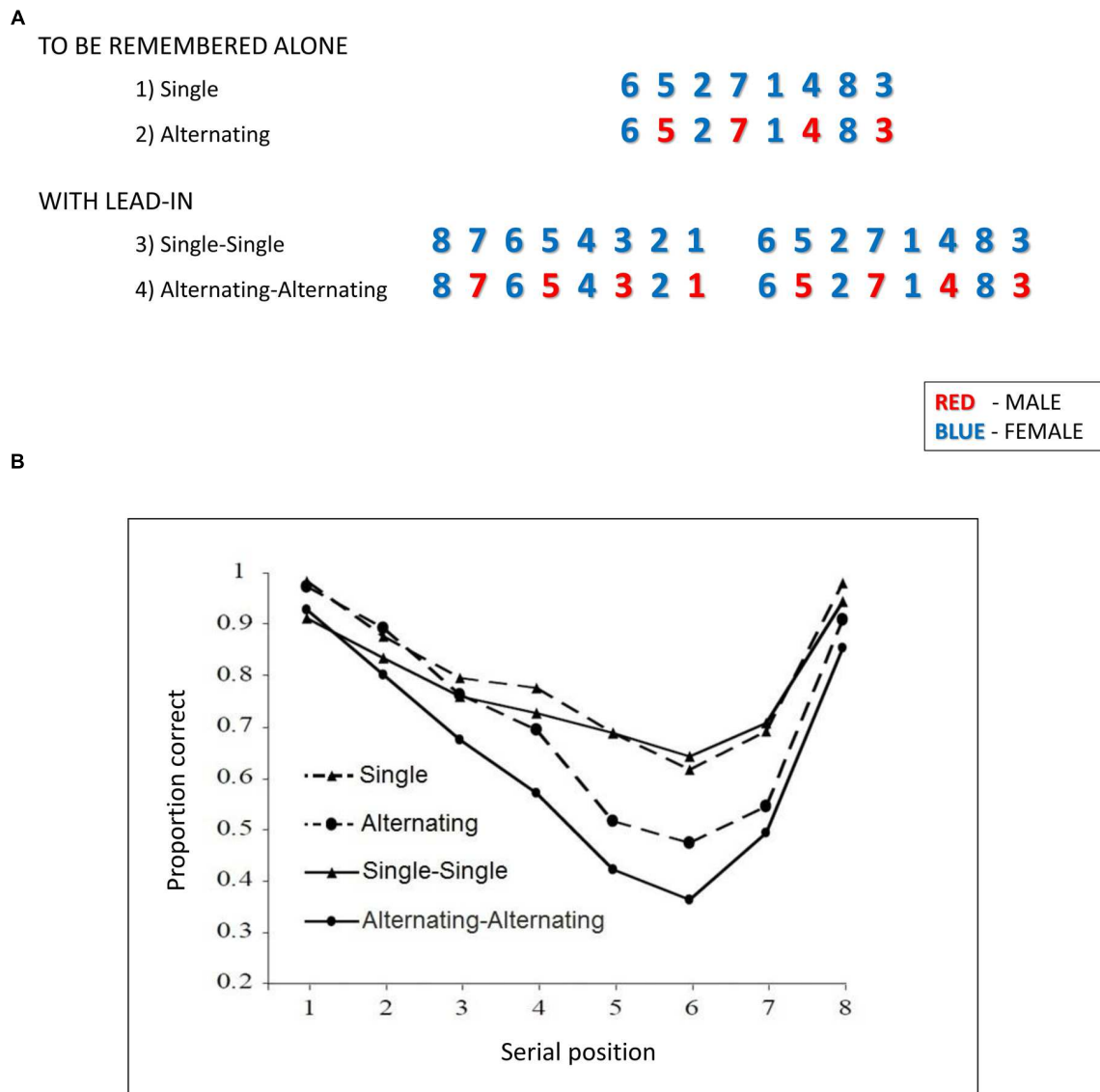
<sup>3</sup>Notably, the mere presence of an irrelevant sequence of repeated items has no appreciable effect on serial recall.



**FIGURE 3 | (A)** Shows schematically the arrangement of stimuli used by Nicholls and Jones (2002). The control condition comprises a sequence of eight to-be-remembered digits. The suffix-only condition has a spoken irrelevant item 'zero' at the end of the to-be-remembered sequence. The Irrelevant + Suffix condition shows a sequence of irrelevant stimuli ('zero') beginning well before the irrelevant sequence and culminating with an item after the last digit in the to-be-remembered sequence. The 'Irrelevant (No Suffix)' condition is the same as the Irrelevant + Suffix condition without a terminal suffix. **(B)** Shows the performance associated with each of those conditions as a function of the presentation position of the stimuli within the to-be-remembered sequence.

item 4). Grouping is effective when it employs exactly those principles of perceptual organization important for reducing the suffix effect. These principles include change of voice, delay in presentation, and change of spatial location, all of which have been confirmed as producing within-list recency effects associated with groups (Frankish, 1989). The principles of grouping in auditory-verbal memory, it appears, are readily inferred from the data showing a reduction of the suffix

effect. Additionally, Frankish (1985) showed that, with *visual* presentation, there is little extra grouping advantage by inserting extra pauses after the third and sixth items in the nine-item list. Frankish (1985) found no obvious difference between the serial position curves produced when participants are asked to subjectively group visual lists and those produced when the presentation of the lists was grouped by half second pauses (Experiment 1).



**FIGURE 4 | (A)** Shows a sub-set of stimuli used by Hughes et al. (2009). To-be-remembered stimuli are first shown in isolation with lists either all from the same voice (Single) and then shown with alternating male and female voices (Alternating). Participants are required to report all the list in the sequence in which it was presented. Then lists with lead-in are shown. In the first case both the lead-in and the to-be-remembered list are in the same voice (Single-Single) and then in alternating voices (Alternating-Alternating). **(B)** Shows the performance associated with each of those conditions as a function of the presentation position of the stimuli within the to-be-remembered sequence.

In a further study, Frankish (1989) showed that an extra pause of only 80 ms following the third and sixth items had as much effect as an extra half-second pause. Likewise, when the middle three digits were differentiated from the others by either voice (male vs. female) or spatial channel (left vs. right ear), the effects of these manipulations were equivalent to those of the temporal change. In addition, the study demonstrated that the voice distinction alone is as effective as voice plus pause. That is, if the middle three digits are in a different voice from the first and last three, then inserting a pause of half a second after the third and sixth digits, thereby, in addition, temporarily isolating the middle three digits, has no further effect.

These effects appear to reflect the automatic segmentation of auditory lists in a manner that is more powerful than the strategic grouping that operates on visually presented lists which produces less of an effect and is more readily disrupted (Hitch et al., 1996). Although a number of researchers (e.g., Hitch et al., 1996; Farrell, 2012) have concentrated on the role of timing—and of extended pauses—in creating groups, Frankish's results clearly show that perceptual groups can be created using cues other than elongated pauses between list items. This observation is important because it shows that factors other than consolidation and rehearsal of a recently completed group (in the pause before the next group arrives) are responsible



for creating these group boundaries. It also shows that the group boundaries can be established very quickly – parsing the list into subgroups almost instantaneously as the stimuli are encountered. Thus, although providing temporal cues to grouping and allowing (or encouraging; Taylor et al., 2015) prosodic, group-based rehearsal to emerge is one means of parsing the input, it is not the only way in which within-list organization can emerge. Crucially for current purposes, the perceptual segmentation of auditory lists is one that requires the constant comparison of the current and preceding auditory input. Automatic overwriting of previous stimuli by incoming information would interfere with the allocation of the current (incoming) stimulus to the appropriate perceptual stream, which may have been established over several preceding items.

## Principles of Organization: Similarity-Based Streaming

Generally, theories of short-term memory fail to acknowledge (or at most, pay lip-service to) the idea that events might be organized—and re-organized—according to perceptual streams. Rather, current theories view short-term memory as post-categorical, item-based encoding within a single, to-be-recalled list. The item here is defined by the experimenter *a priori* rather than inferred from the behavior of the participant. Those characteristics of the stimuli that denote common origin, that connote streams—among them similarity of pitch, timbre, location, and proximity in time—are ignored by such accounts, which also overlook the fluidity and flexibility of systems within which items are organized—and re-organized—according to their perceived belonging to one or more sources of origin. We argue that this is a profound mistake.

In the first instance, it is logical to assume that whatever form representations take in memory is constrained by the way in which information is available perceptually. The existence of natural organizational principles, known since the advent of Gestalt psychology, implies that multiple streams of information co-exist within memory in a way that is inconsistent with strict overwriting as the mechanism for forgetting. In the second instance, treating memoranda as discrete and independent items within the experimental participants' cognitive systems because they were conceived and presented as such by the experimenter is an unwarranted assumption. The assumption arises directly from the idea that representations are, almost by definition, abstract and “post-categorical,” whereas in fact very few studies have examined the extent to which memory results can be accounted for by categorical vs. *continuous* storage systems (Frankish, 2008; Joseph et al., 2015). Taken to the extreme, it is clear that the recall of individual items is not independent, and whilst few models make this mistake, the amount and type of information relating the experimenter-defined items to one another and to a perceived locus of origin is impoverished in current theories. The relationship between items is formally often one merely of time or position (e.g., Page and Norris, 1998; Brown et al.,

2007). Commonality of perceptual characteristics rarely plays a role because all of the elements within the memoranda are automatically assigned to a single list-structure, something that presumably occurs at a pre-mnemonic processing stage. Where between-item similarity is considered (as for example, to model the phonological similarity effect) this may often be at a distinct stage from positional similarity. For example, the primacy model of Page and Norris (1998) in which positional errors between localist representations occur naturally along the “primacy gradient” then forward items onto an explicitly phonological distributed representation stage prior to output in order to implement item confusion errors (Beaman, 2000)<sup>4</sup>.

Missing from all of these accounts is any measure of *stream-based similarity* such that elements within the memoranda are allocated to one stream or another based upon a common theme or thread running through the sequence and which serves to distinguish this stream from another. Stream-based similarity, according to this analysis, is necessary to account for the effects reviewed above – the reduction or elimination of suffix effects, the talker variability effect, the perceptual grouping effect and so on. The thread of similarity that acts to hold elements together is, however, precisely the source of interference that would consistently and continually degrade individual item representations under an overwriting account.

The availability of information about the stream to which the stimuli belong is precisely what is needed to account for moderation and abolition of suffix effects, between-talker variability effects, and within-list grouping effects as reviewed here. Discontinuities in time (i.e., elongated breaks between groups) have been used to account for within-list temporal grouping effects (Nairne, 1990; Hitch et al., 1996; Farrell, 2012). This mechanism follows naturally from the idea of overwriting, since a break is naturally interpreted as a pause in which information can be consolidated and/or within which retroactive interference (such as overwriting) will not occur. Such accounts do not properly address the effects of very short pauses between groups which are more parsimoniously conceived of as groupings caused by discontinuities in rhythm rather than time *per se*, nor are they able to account for grouping effects caused by intonation, timbre or spatial location. For the same reasons, speaker-variability effects and reduced suffix effects are not predicted by such accounts because the models do not maintain the correct types of information to give rise to such effects. To do so, not only must information about physical characteristics be maintained in addition to whatever post-categorical or more abstract labels that may be assumed, but also information must be held about the stream as a whole rather than individual items in isolation, and incoming information (e.g., a post-list suffix) interpreted in terms of the information held and

<sup>4</sup>This is a mirror image of how the feature model addresses the same situation: in the feature model, item-based confusions arise naturally from the distributed representation of items as vectors of feature values but positional errors only occur when an item independently “drifts” along the position dimension (Neath, 1999, 2000).

prior expectations it elicits, as shown both by contextual suffix effects and by experiments repeating and streaming the suffix. The main conclusions point to the intimacy of perception and memory, or perhaps even to their wholesale integration. Certainly, no attribution to the action of auditory memory should be entertained until a thorough analysis of how auditory streaming could explain the same phenomena has been dismissed. Only after streaming processes have yielded the super-ordinate structure of the material being remembered can other approaches – such as overwriting – be entertained as explanatory constructs.

## REFERENCES

- Altmann, E. M. (2009). Evidence for temporal decay in short-term episodic memory. *Trends Cogn. Sci.* 13, 279–279. doi: 10.1016/j.tics.2009.04.001
- Altmann, E. M., and Gray, W. D. (2002). Forgetting to remember: the functional relationship of decay and interference. *Psychol. Sci.* 13, 27–33. doi: 10.1111/1467-9280.00405
- Altmann, E. M., and Schunn, C. D. (2012). Decay versus Interference: a new look at an old interaction. *Psychol. Sci.* 23, 1435–1437. doi: 10.1177/0956797612446027
- Ayres, T. J., Jonides, J., Reitman, J. S., Egan, J. C., and Howard, D. A. (1979). Differing suffix effects for the same physical suffix. *J. Exp. Psychol. Hum. Learn. Mem.* 5, 315–321.
- Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Q. J. Exp. Psychol.* 18, 302–309. doi: 10.1080/14640746608400055
- Barrouillet, P., De Pape, A., and Langerock, N. (2012). Time causes forgetting from working memory. *Psychon. Bull. Rev.* 19, 87–92. doi: 10.3758/s13423-011-0192-8
- Barrouillet, P., Portrat, S., Vergauwe, E., Diependaele, K., and Camos, V. (2011). Further evidence for temporal decay in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 1302–1317.
- Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press.
- Beaman, C. P. (2000). Neurons amongst the symbols? *Behav. Brain Sci.* 23, 468–470. doi: 10.1017/S0140525X00233359
- Beaman, C. P. (2002). Inverting the modality effect in serial recall. *Q. J. Exp. Psychol.* 55A, 371–389. doi: 10.1080/02724980143000307
- Beaman, C. P., and Jones, D. M. (1998). Irrelevant sound disrupts order information in free recall as in serial recall. *Q. J. Exp. Psychol.* 51A, 615–636. doi: 10.1080/713755774
- Beaman, C. P., Neath, I., and Surprenant, A. M. (2008). Modeling distributions of immediate memory effects: no strategies needed? *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 219–229. doi: 10.1037/0278-7393.34.1.219
- Berman, M. G., Jonides, J., and Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Mem. Cogn.* 35, 317–333. doi: 10.1037/a0014873
- Bhatara, P., Ward, G., and Tan, L. (2006). Examining the relationship between free recall and immediate serial recall: the effect of concurrent task performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 215–229.
- Bhatara, P., Ward, G., and Tan, L. (2008). Examining the relationship between free recall and immediate serial recall: the serial nature of recall and the effect of test expectancy. *Mem. Cogn.* 36, 20–34. doi: 10.3758/MC.36.1.20
- Botvinick, M., and Plaut, D. C. (2006). Short-term memory for serial order: a recurrent neural network model. *Psychol. Rev.* 113, 201–233. doi: 10.1037/0033-295X.113.2.201
- Bregman, A. S. (1990). *Auditory Scene Analysis*. Cambridge, MA: MIT Press.
- Bregman, A. S., and Rudnick, A. I. (1975). Auditory segregation: stream or streams? *J. Exp. Psychol. Hum. Percept. Perform.* 1, 263–267.
- Broadbent, D. E. (1958). *Perception and Communication*. London: Pergamon.
- Brown, G. D., Preece, T., and Hulme, C. (2000). Oscillator-based memory for serial order. *Psychol. Rev.* 107, 127–181. doi: 10.1037/0033-295X.107.1.127
- Brown, G. D. A., and Lewandowsky, S. (2010). “Forgetting in memory models: arguments against trace decay and consolidation failure,” in *Forgetting*, ed. S. Della Sala (Hove: Psychology Press), 49–75.
- Brown, G. D. A., Neath, I., and Chater, N. (2007). A temporal ratio model of memory. *Psychol. Rev.* 114, 539–576. doi: 10.1037/0033-295X.114.3.539
- Burgess, N., and Hitch, G. J. (1999). Memory for serial order: a network model of the phonological loop and its timing. *Psychol. Rev.* 106, 551–581. doi: 10.1037/0033-295X.106.3.551
- Burgess, N., and Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *J. Mem. Lang.* 55, 627–652. doi: 10.1016/j.jml.2006.08.005
- Carr, D., and Miles, C. (1997). Rhyme attenuates the auditory suffix effect: alliteration does not. *Q. J. Exp. Psychol.* 50A, 518–527. doi: 10.1080/027249897392008
- Conrad, R. (1960). Serial order intrusions in immediate memory. *Br. J. Psychol.* 51, 45–48. doi: 10.1111/j.2044-8295.1960.tb00723.x
- Conrad, R. (1964). Acoustic confusions in immediate memory. *Br. J. Psychol.* 55, 75–84. doi: 10.1111/j.2044-8295.1964.tb00899.x
- Conrad, R., and Hull, A. J. (1964). Information, acoustic confusion, and memory span. *Br. J. Psychol.* 55, 429–432. doi: 10.1111/j.2044-8295.1964.tb00928.x
- Craik, F. I. M., and Watkins, M. J. (1973). The role of rehearsal in short-term memory. *J. Verb. Learn. Verb. Behav.* 12, 599–607. doi: 10.1016/S0022-5371(73)80039-8
- Crowder, R. G. (1971). Waiting for the stimulus suffix: decay, delay, rhythm, and readout in immediate memory. *Q. J. Exp. Psychol.* 23, 324–340. doi: 10.1080/14640746908401829
- Crowder, R. G. (1978). Mechanisms of auditory backward masking in the stimulus suffix effect. *Psychol. Rev.* 85, 502–524. doi: 10.1037/0033-295X.85.6.502
- Crowder, R. G., and Cheng, C.-M. (1973). Phonemic confusability, precategorical acoustic storage, and the suffix effect. *Percept. Psychophys.* 13, 145–148. doi: 10.3758/BF03207250
- Crowder, R. G., and Morton, J. (1969). Precategorical acoustic storage (PAS). *Percept. Psychophys.* 5, 365–373. doi: 10.3758/BF03210660
- Deutsch, D. (1972). Effect of repetition of standard and comparison tones on recognition memory for pitch. *J. Exp. Psychol.* 93, 156–162. doi: 10.1037/h0032496
- Deutsch, D. (1978a). Delayed pitch comparisons and the principle of proximity. *Percept. Psychophys.* 23, 227–230. doi: 10.3758/BF03204130
- Deutsch, D. (1978b). Interference in pitch memory as a function of ear of input. *Q. J. Exp. Psychol.* 30, 282–287. doi: 10.1080/14640747808400675
- Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychol. Rev.* 119, 223–271. doi: 10.1037/a0027371
- Fournet, N., Juphard, A., Monnier, C., and Roulin, J.-L. (2003). Phonological similarity in free and serial recall: the effect of increasing retention intervals. *Int. J. Psychol.* 38, 384–389. doi: 10.1080/00207590344000204
- Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *J. Mem. Lang.* 24, 200–209. doi: 10.1016/j.cognition.2013.08.017
- Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 469–479.
- Frankish, C. (1995). Intonation and auditory grouping in immediate serial recall. *Appl. Cogn. Psychol.* 9, 5–22. doi: 10.1002/acp.2350090703

## AUTHOR CONTRIBUTIONS

This manuscript was co-written by CPB and DJ. Figures were provided by DJ.

## FUNDING

The research reported in this article was supported by Economic and Social Research Council (UK) grant ES/L00710X/1 awarded to Philip Beaman and Dylan M. Jones.

- Frankish, C. (2008). Precategorical acoustic storage and the perception of speech. *J. Mem. Lang.* 58, 815–836. doi: 10.1016/j.jml.2007.06.003
- Frankish, C., and Turner, J. (1984). Delayed suffix effect at very short delays. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 767–777.
- Grenfell-Essam, R., and Ward, G. (2012). Examining the relationship between free recall and immediate serial recall: the role of list length, strategy use, and test expectancy. *J. Mem. Lang.* 67, 106–148. doi: 10.1016/j.jml.2012.04.004
- Grenfell-Essam, R., Ward, G., and Tan, L. (2013). The role of rehearsal on the output order of immediate free recall of short and long lists. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 317–347. doi: 10.1037/a0028974
- Henson, R. N. A. (1998). Short-term memory for serial order: the start-end model. *Cogn. Psychol.* 36, 73–137. doi: 10.1006/cogp.1998.0685
- Hitch, G. J., Burgess, N., Towse, J. N., and Culpin, V. (1996). Temporal grouping effects in immediate recall: a working memory analysis. *Q. J. Exp. Psychol.* 49A, 116–139. doi: 10.1080/027249896392829
- Hughes, R. W., Chamberland, C., Tremblay, S., and Jones, D. M. (in press). Perceptual-motor determinants of auditory-verbal serial short-term memory. *Journal of Memory and Language*.
- Hughes, R. W., Marsh, J. E., and Jones, D. M. (2009). Perceptual-gestural (mis)mapping in serial short-term memory: the impact of talker variability. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 1411–1425. doi: 10.1037/a0017008
- Jones, D. M. (1993). “Objects, streams and threads of auditory attention,” in *Attention: Selection, Awareness, and Control: A Tribute to Donald Broadbent*, eds A. D. Baddeley and L. Weiskrantz (New York, NY: Clarendon Press), 87–104.
- Jones, D. M. (1999). The cognitive psychology of auditory distraction: the 1997 BPS Broadbent Lecture. *Br. J. Psychol.* 90, 167–187. doi: 10.1348/000712699161314
- Jones, D. M., Alford, D., Bridges, A., Tremblay, S., and Macken, W. J. (1999). Organizational factors in selective attention: the interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 464–473.
- Jones, D. M., Beaman, C. P., and Macken, W. J. (1996). “The object-oriented episodic record model,” in *Models of Short-Term Memory*, ed. S. E. Gathercole (Hove: Psychology Press), 209–238.
- Jones, D. M., Macken, W. J., and Harries, C. (1997). Disruption of short-term recognition memory for tones: streaming or interference? *Q. J. Exp. Psychol. Learn. Mem. Cogn.* 50A, 337–357. doi: 10.1080/713755707
- Joseph, S., Iverson, P., Manohar, S., Fox, Z., Scott, S. K., and Husain, M. (2015). Precision of working memory for speech sounds. *Q. J. Exp. Psychol. Learn. Mem. Cogn.* 68, 2022–2040. doi: 10.1080/17470218.2014.1002799
- Lange, E., and Oberauer, K. (2005). Overwriting of phonemic features in serial recall. *Memory* 13, 333–339. doi: 10.1080/09658210344000378
- Lashley, K. S. (1951). “The problem of serial order in behavior,” in *Cerebral Mechanisms in Behavior. The Hixon Symposium*, ed. L. A. Jeffress (New York, NY: Wiley), 112–136.
- Lewandowsky, S., and Farrell, S. (2011). *Computational Modelling in Cognition: Principles and Practice*. Thousand Oaks, CA: Sage.
- Lewandowsky, S., Geiger, S. M., and Oberauer, K. (2008). Interference-based forgetting in verbal short-term memory. *J. Mem. Lang.* 59, 200–222. doi: 10.1016/j.tics.2008.12.003
- Lewandowsky, S., and Murdock, B. B. Jr. (1989). Memory for serial order. *Psychol. Rev.* 96, 25–57. doi: 10.1037/0033-295X.96.1.25
- Lewandowsky, S., and Oberauer, K. (2008). The word length effect provides no evidence for decay in short-term memory. *Psychon. Bull. Rev.* 15, 875–888. doi: 10.3758/PBR.15.5.875
- Lewandowsky, S., and Oberauer, K. (2009). No evidence for temporal decay in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 1545–1551. doi: 10.1037/a0017010
- Lewandowsky, S., and Oberauer, K. (2015). Rehearsal in serial recall: an unworkable solution to the non-existent problem of decay. *Psychol. Rev.* 122, 674–699. doi: 10.1037/a0039684
- Lewandowsky, S., Oberauer, K., and Brown, G. D. A. (2009). No temporal decay in verbal short-term memory. *Trends Cogn. Sci.* 13, 120–126. doi: 10.1016/j.tics.2008.12.003
- Macken, B., Taylor, J., and Jones, D. (2015). Limitless capacity: a dynamic object-oriented approach to short-term memory. *Front. Psychol.* 6:293. doi: 10.3389/fpsyg.2015.00293
- Mathias, S. R., and von Kriegstein, K. (2014). Percepts, not acoustic properties, are the units of auditory short-term memory. *J. Exp. Psychol. Learn. Mem. Cogn. Hum. Percept. Perform.* 40, 445–450.
- Mercer, T., and McKeown, D. (2010a). Interference in short-term auditory memory. *Q. J. Exp. Psychol. Learn. Mem. Cogn.* 63, 1256–1265. doi: 10.1080/17470211003802467
- Mercer, T., and McKeown, D. (2010b). Updating and feature overwriting in short-term memory for timbre. *Atten. Percept. Psychophys.* 72, 2289–2303. doi: 10.3758/APP.72.8.2289
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 101, 343–352. doi: 10.1037/0033-295X.101.2.343
- Morton, J. (1964). A preliminary functional model for language behaviour. *Int. Audiol.* 3, 216–225. doi: 10.3109/05384916409074089
- Morton, J. (1969). Interaction of information in word recognition. *Psychol. Rev.* 76, 165–178. doi: 10.1037/h0027366
- Morton, J. (1976). Two mechanisms in the stimulus suffix effect. *Mem. Cogn.* 4, 144–149. doi: 10.3758/BF03213156
- Morton, J. (1979). “Facilitation in word recognition: experiments causing change in the logogen model,” in *Processing of Visible Language*, eds P. A. Kollers, M. E. Wroldstad, and H. Bouma (New York, NY: Plenum), 259–268.
- Morton, J., and Chambers, S. M. (1976). Some evidence for “speech” as an acoustic feature. *Br. J. Psychol.* 67, 31–45. doi: 10.1111/j.2044-8295.1976.tb01495.x
- Morton, J., Crowder, R. G., and Prussin, H. A. (1971). Experiments with the stimulus suffix effect. *J. Exp. Psychol.* 91, 169–190. doi: 10.1037/h0031844
- Morton, J., Marcus, S. M., and Ottley, P. (1981). The acoustic correlates of “speechlike”: a use of the suffix effect. *J. Exp. Psychol. Learn. Mem. Cogn. Gen.* 110, 568–593.
- Murdock, B. B. Jr. (1968). Serial order effects in short-term memory. *J. Exp. Psychol. Learn. Mem. Cogn. Monogr. Suppl.* 76, 1–15. doi: 10.1037/h0025694
- Murdock, B. B. Jr. (1983). A distributed model for serial-order information. *Psychol. Rev.* 90, 316–338. doi: 10.1037/0033-295X.90.4.316
- Nairne, J. S. (1990). A feature model of immediate memory. *Mem. Cogn.* 18, 251–269. doi: 10.3758/BF03213879
- Nairne, J. S. (2002). Remembering over the short-term: the case against the standard model. *Annu. Rev. Psychol.* 53, 53–81. doi: 10.1146/annurev.psych.53.100901.135131
- Nairne, J. S., and Kelley, M. R. (1999). Reversing the phonological similarity effect. *Mem. Cogn.* 27, 45–53. doi: 10.3758/BF03201212
- Neath, I. (1999). Modelling the disruptive effects of irrelevant speech on order information. *Int. J. Psychol.* 34, 410–418. doi: 10.1080/002075999399765
- Neath, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychon. Bull. Rev.* 7, 403–423. doi: 10.3758/BF03214356
- Neath, I. (2010). Evidence for similar principles in episodic and semantic memory: the presidential serial position function. *Mem. Cogn.* 38, 659–666. doi: 10.3758/MC.38.5.659
- Neath, I., and Brown, G. D. A. (2012). Arguments against memory trace decay: a SIMPLE account of Baddeley & Scott. *Front. Cogn.* 3:35. doi: 10.3389/fpsyg.2012.00035
- Neath, I., and Nairne, J. S. (1995). Word-length effects in immediate memory: overwriting trace-decay theory. *Psychon. Bull. Rev.* 2, 429–441. doi: 10.3758/BF03210981
- Neath, I., Surprenant, A. M., and Crowder, R. G. (1993). The context-dependent stimulus suffix effect. *J. Exp. Psychol. Learn. Mem. Cogn. Learn. Mem. Cogn.* 19, 698–703.
- Nicholls, A. P., and Jones, D. M. (2002). Capturing the suffix: cognitive streaming in immediate serial recall. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 12–28.
- Oberauer, K. (2009). Interference between storage and processing in working memory: feature overwriting, not similarity-based competition. *Mem. Cogn.* 37, 346–357. doi: 10.3758/MC.37.3.346
- Oberauer, K., and Kliegl, R. (2006). A formal model of capacity limits in working memory. *J. Mem. Lang.* 55, 601–626. doi: 10.1037/a0025660
- Oberauer, K., and Lange, E. B. (2008). Interference in verbal working memory: distinguishing similarity-based confusion, feature overwriting, and feature migration. *J. Mem. Lang.* 58, 730–745. doi: 10.1016/j.jml.2007.09.006



- Oberauer, K., and Lewandowsky, S. (2008). Forgetting in immediate serial recall: decay, temporal distinctiveness, or interference? *Psychol. Rev.* 115, 544–576. doi: 10.1037/0033-295X.115.3.544
- Oberauer, K., and Lewandowsky, S. (2013). Evidence against decay in verbal working memory. *J. Exp. Psychol. Learn. Mem. Cogn. Gen.* 142, 380–411. doi: 10.1037/a0029588
- Ottley, P., Marcus, S., and Morton, J. (1982). Contextual effects in the stimulus suffix paradigm. *Br. J. Psychol.* 73, 383–387. doi: 10.1111/j.2044-8295.1982.tb01820.x
- Page, M. P. A., and Norris, D. (1998). The primacy model: a new model for serial recall. *Psychol. Rev.* 105, 761–781. doi: 10.1037/0033-295X.105.4.761-781
- Portrat, S., Barrouillet, P., and Camos, V. (2008). Time-related decay or interference-based forgetting in working memory? *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 1561–1564.
- Raffone, A., and Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *J. Cogn. Neurosci.* 13, 766–785. doi: 10.1162/08989290152541430
- Ricker, T. J., Vergauwe, E., and Cowan, N. (2014). Decay theory of immediate memory: from Brown (1958) to today (2014). *Q. J. Exp. Psychol.* doi: 10.1080/17470218.2014.914546 [Epub ahead of print],
- Saint-Aubin, J., and Poirier, M. (1999). Semantic similarity and immediate serial recall: is there a detrimental effect on order information? *Q. J. Exp. Psychol. Learn. Mem. Cogn.* 52A, 367–394. doi: 10.1080/713755814
- Scharf, B. (1978). "Loudness," in *Handbook of Perception*, Vol. 4, eds E. C. Carterette and M. P. Friedman (New York, NY: Academic Press), 187–242.
- Semal, C., and Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. *J. Acoust. Soc. Am.* 89, 2404–2410. doi: 10.1121/1.400928
- Semal, C., and Demany, L. (1993). Further evidence for an autonomous processing of pitch in auditory short-term memory. *J. Acoust. Soc. Am.* 94, 1315–1322. doi: 10.1121/1.408159
- Shallice, T., and Cooper, R. (2010). *The Organisation of Mind*. Oxford: Oxford University Press.
- Spurgeon, J., Ward, G., and Matthews, W. (2014). Examining the relationship between immediate serial recall and immediate free recall: common effects of Phonological Loop variables, but only limited evidence for the Phonological Loop. *J. Exp. Psychol. Learn. Mem. Cogn. Learn. Mem. Cogn.* 40, 1110–1141. doi: 10.1037/a0035784
- Starr, G. E., and Pitt, M. (1997). Interference effects in short-term memory for timbre. *J. Acoust. Soc. Am.* 102, 486–494. doi: 10.1121/1.419722
- Surprenant, A. M., and Neath, I. (2009). *Principles of Memory*. New York, NY: Psychology Press.
- Taylor, J. C., Macken, W. J., and Jones, D. M. (2015). A matter of emphasis: linguistic stress habits modulate serial recall. *Mem. Cogn.* 43, 520–537. doi: 10.3758/s13421-014-0466-2
- Ward, G., and Tan, L. (2004). The effect of the length of to-be-remembered lists and intervening lists on free recall: are-examination using overt rehearsal. *J. Exp. Psychol. Learn. Mem. Cogn. Learn. Mem. Cogn.* 30, 1196–1210.
- Ward, G., Tan, L., and Grenfell-Essam, R. (2010). Examining the relationship between free recall and immediate serial recall: the effects of list length and output order. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 1207–1241. doi: 10.1037/a0020122
- Ward, G., Woodward, G., Stevens, A., and Stinson, C. (2003). Using overt rehearsals to explain word frequency effects in free recall. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 186–210.
- Warren, R. M., Obusek, C. J., Farmer, R. M., and Warren, R. P. (1969). Auditory sequence: confusion of patterns other than speech or music. *Science* 164, 586–587. doi: 10.1126/science.164.3879.586
- Watkins, M. J., Watkins, O. C., and Crowder, R. G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *J. Verb. Learn. Verb. Behav.* 13, 430–447. doi: 10.1016/S0022-5371(74)80021-6

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Beaman and Jones. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.